WEB PHISHING DETECTION

LITERATURE SURVEY

TITLE : Detection of Phishing Websites through Computational Intelligence

Publisher: IEEE

Year: 2021

DESCRIPTION:

Phishing is a technique used by hackers to fool internet users reveal their sensitive information like passwords, credit card numbers, contact information, and address, etc. Web phishing is carried out mostly by sending fake web links to the users through different communication means like Email, Facebook Messenger and WhatsApp, etc. Web phishing detection is significant for making internet browsing safe and secure for users. Different approaches were applied for the detection of fake websites. However, the most efficient method for detecting phishing websites is the one that is based on artificial intelligence and learning mechanism. In this research, an efficient and accurate method is proposed for the detection of phishing websites which is based on computational intelligence. Through the development of different computational models and rigorous testing, it was revealed that Extreme Gradient Boost (XGBoost) based model achieved the maximum scores in all the validation tests. This shows that the model is robust and accurate in terms of web-phishing detection.

TITLE: Phishing Website Detection Based on URL

Author: Salvi Siddhi Ravindra , Shah Juhi Sanjay, Shaikh Nausheenbanu Ahmed Gulzar, Khodke Pallavi

Year:2021

DESCRIPTION:

In today's era, due to the surge in the usage of the internet and other online platforms, security has been major attention. Many cyberattacks take place each day out of which website phishing is the most common issue. It is an

act of imitating a legitimate website and thereby tricking the users and stealing their sensitive information. So, concerning this problem, this paper will introduce a possible solution to avoid such attacks by checking whether the provided URLs are phishing URLs or legitimate URLs. It is a Machine Learning based system especially Supervised learning where we have provided 2000 phishing and 2000 legitimate URL dataset. We have taken into consideration the Random Forest Algorithm due to its performance and accuracy. It considers 9 features and hence detects whether the URL is safe to access or a phishing URL. Keywords : URLs, Phishing, Legitimate, Machine Learning.

TITLE: Detection of Phishing Websites using Machine Learning

Author: Atharva Deshpande, Omkar Pedamkar , Nachiket, Dr. Swapna Borde

Year:2021

DESCRIPTION:

Phishing is a common attack on credulous people by making them to disclose their unique information using counterfeit websites. The objective of phishing website URLs is to purloin the personal information like user name, passwords and online banking transactions. Phishers use the websites which are visually and semantically similar to those real websites. As technology continues to grow, phishing techniques started to progress rapidly and this needs to be prevented by using anti-phishing mechanisms to detect phishing. Machine learning is a powerful tool used to strive against phishing attacks. This paper surveys the features used for detection and detection techniques using machine learning. Phishing is popular among attackers, since it is easier to trick someone into clicking a malicious link which seems legitimate than trying to break through a computer’s defense systems. The malicious links within the body of the message are designed to make it appear that they go to the spoofed organization using that organization’s logos and other legitimate contents. Here, we explain phishing domain (or Fraudulent Domain) characteristics, the features that distinguish them from legitimate domains, why it is important to detect these domains, and how they can be detected using machine learning and natural language processing techniques.

TITLE: Detection of Phishing Websites using an Efficient Machine Learning Framework

Author: Naresh Kumar D, Premnath, Nishanth Kumar, Nemala Sai Rama Hemanth

Year:2020

DESCRIPTION:

Phishing attack is one of the commonly known attack where the information from the internet users are stolen by the intruder. The internet users are losses their sensitive information such as Protected passwords, personal information and their transactions to the intruders. The Phishing attack is normally carried by the attackers where the legitimate frequently used websites are manipulated and masked to gather the personal information of the users. The Intruders use the personal information and can manipulate the transactions and get definite from them. From the literature there are various anti-Phishing websites by the various authors. Some of the techniques are Blacklist or Whitelist and heuristic and visual similarity based methods. In spite of the users using these techniques most of the users are getting attacked by the intruders by means of Phishing to gather their sensitive information. A novel Machine Learning based classification algorithm has been proposed in this paper which uses heuristic features where feature selection can be extracted from the attributes such as Uniform Resource Locator, Source Code, Session, Type of security involve, Protocol used, type of website. The proposed model has been evaluated using five machine learning algorithms such as random forest, K Nearest Neighbor, Decision Tree, Support Vector Machine, Logistic regression. Out of these models, the random forest algorithm performs better with attack detection accuracy of 91.4%. Moreover the Random Forest Model uses orthogonal and oblique classifiers to select the best classifiers for accurate detection of Phishing attacks in the websites.

TITLE: A Review Paper on Detection of Phishing Websites using Machine Learning

Author: Ashritha Jain R, Mrs. Mangala Kini, Chaithra Kulal, Deekshitha

Year: 2020

DESCRIPTION:

Phishing is the fraudulent attempt to obtain sensitive information of individuals or organization such as usernames, passwords and credit card details by disguising as trustworthy entity in a electronic communication. Phishing attack causes serious threats to user’s privacy and security. The purpose of this study is to presents an overview about various phishing attacks and various techniques to protect the information. It also includes the discussion of Extreme Learning Machine (ELM) based classification for 30 features including phishing websites data in UC Irvine Machine Learning Repository database. Keywords- Phishing, Extreme Learning Machine.